ELSEVIER

International Journal of Industrial Ergonomics 26 (2000) 277–283

International Journal of Industrial Ergonomics

www.elsevier.nl/locate/ergon

Predicting the accuracy of visual search performance in the structural inspection of aircraft

Brian J. Melloy*, John M. Harris, Anand K. Gramopadhye

Department of Industrial Engineering, College of Engineering and Science, Clemson University, Clemson, SC 29634-0920, USA

Abstract

The predominant tasks in aircraft inspection are those that require visual search. Speed and accuracy characterize visual search tasks. Since the trade-off between speed and accuracy directly affects the safety, dependability and affordability of air transportation, there is considerable motivation to express this relationship in quantitative terms. Although models of this trade-off have been previously proposed for various search tasks, the applicability of these models to the tasks typically required of aircraft inspectors is limited. Thus new models of visual search that typify aircraft inspection are adopted here in order to examine the explicit trade-off between speed and accuracy in this environment.

Relevance to industry

The models adopted here, specifically applicable to aircraft inspection, consider the trade-off between speed and accuracy – a trade-off that ultimately affects the affordability, dependability, and safety (and therefore public perception of) air travel – in quantitative terms. These models also define the extent to which accuracy can be improved.

Keywords: Inspection; Visual search; Speed and accuracy

1. Introduction

The goal of an aviation inspection and maintenance system is to provide the public with safe, dependable, and affordable air transportation. Accordingly, the aviation industry intensified its efforts to examine its inspection procedures in response to a number of maintenance-related aircraft accidents and incidents in the late-1980s (FAA, 1991). The investigation of inspection procedures was warranted by the fact that approximately 18% of these wide-body aircraft accidents were a consequence of maintenance and inspection errors, the majority of which were attributed to inspection (Phillips, 1994).

Simply stated, the inspection function involves examining structures for defects that can affect the airworthiness of the aircraft. The primary structural defects that occur on structures are cracks and corrosion (FAA, 1991); cracks are a result of repeated stretching of the structure from aerodynamic or internal pressure loads, while corrosion arises

* Corresponding author. Tel.: + 1-864-656-5540; fax: + 1-864-656-0795.

E-mail address: brian.melloy@ces.clemson.edu (B.J. Melloy).

0169-8141/00/$ - see front matter
PII: S0169-8141(99)00071-2

from weathering or exposure to harmful chemicals. Older aircraft are more susceptible to the effects of fatigue cracks (especially multi-site damage) and corrosion (Drury et al., 1990). Since almost the entire commercial fleet is now operating into the "extended life" phase of its life cycle (Bobo, 1990), the inspection activity has become even more critical.

Despite the fact that aircraft inspection includes a number of diverse tasks (e.g., nondestructive tests, such as employing eddy current equipment to detect cracks in rivets; tactile tasks, such as wiping hoses in order to detect fluid leaks; and visual tasks, such as searching for corrosion on aircraft structures), studies of the process have revealed that visual search constitutes 90% of these tasks (Drury et al., 1990). Accordingly, visual search is the subject of this investigation. (It should be emphasized that the visual search component of the inspection task is distinct from the decision making component. Moreover, it is not unusual for the mechanics, rather than the inspectors, to make the decisions regarding the nature and severity of the defects.)

Speed and accuracy characterize the performance of visual search tasks. Speed refers to the time required to complete the task, whereas accuracy is related to the number of defects detected. These two measures are inversely related; that is, accuracy generally decreases as speed increases and vice versa. This relationship, which has been validated both in laboratory settings and under field conditions, is commonly referred to as the speed accuracy trade-off (SATO) (Drury, 1994).

The trade-off between speed and accuracy affects the safety, dependability, and affordability of air transportation. For example, if the time constraint were to be relaxed, accuracy, and therefore safety, would be expected to improve as a result. Moreover, the number of inspection/maintenance/inspection[1] cycles could be reduced, which would be desirable from a logistical standpoint. On the other hand, such a policy could conceivably disrupt flight schedules. It would certainly be more costly to an industry that spent over eight billion dollars for maintenance in calendar year 1988 (FAA, 1991).

Thus it is not only clear that speed and accuracy affect the objectives of safety, dependability, and affordability, but that these objectives can also be conflicting. These circumstances provide compelling motivation to analyze the SATO in quantitative terms. Such an analysis though, requires models of visual search that are capable of satisfactorily predicting inspector accuracy as a function of time in this domain.

Several different models have been employed previously to investigate the SATO (e.g., Morawski et al., 1980,1992; Arani et al., 1984; Karwan et al., 1995; Drury and Chi, 1995). These models, however, were designed expressly for two situations: (1) situations in which only one defect of a specific type could occur (e.g., Morawski et al., 1992; Drury and Chi, 1995), or (2) situations wherein multiple defects could occur, but with the stipulation that the search would terminate when one or more defects are detected (e.g., Morawski et al., 1980; Arani et al., 1984). Moreover in these instances, an item is ordinarily classified as "defective" if even one defect is detected in the search field. Consequently, search accuracy has traditionally been defined as the proportion of defective items that are discovered.

In contrast, the models later developed by Harris et al. (1998) were designed for situations in which the objective is to locate as many defects as possible on a part, within a specified period of time. Accordingly, accuracy is defined in this case as the proportion of defects that are discovered in the search field. Since the design of these models and the associated measure of accuracy are more consistent with that of aircraft inspection tasks, there is cause to reexamine this issue. Therefore, the models developed by Harris et al. (1998) will be employed to analyze the relationship between speed and accuracy in this specific context.

2. Model characterization

The process of searching a field (e.g., an aircraft structure) for defects is modeled as a series of fixations. The search field itself is represented as a set of uniformly sized cells. The size of these cells

[1] These latter follow-up inspections are often referred to as buy-back inspections.

correspond to the *visual lobe*, or, in other words, "the area … which can be perceived in a single glimpse" or fixation (Morawski et al., 1992). Any one of these cells may contain one or more defects. In order for a particular defect to be located, two events must occur in succession. First, the inspector must first fixate on the cell that contains the defect and secondly, the inspector must detect the defect.

Whether or not an inspector fixates on a particular cell depends on the search behavior and the number of fixations (i.e., the time engaged in search). Systematic and random patterns of behavior are exhibited throughout the search process. An inspector displaying systematic search behavior will choose from among the cells that have not yet been fixated on, whereas the subsequent fixations of an inspector exhibiting random behavior will be arbitrary. (Search behavior is commonly assumed to be influenced by memory retrieval (e.g., Arani, 1981) and search strategy (e.g., Williams, 1966; Drury and Chi, 1995).) Thus, these two tendencies represent behavioral extremes.

The likelihood of an inspector fixating on a particular cell is directly related to the number of fixations, relative to the size of the search field, for any established search behavior. Of particular interest here is the number of fixations required to make a complete scan (exhaustive search) of the field, since failing to do so would increase the risk of an inspector overlooking defects. (This is in contrast to traditional applications in which a search would terminate upon the detection of any defect.) However, an exhaustive scan of the field can only be assured in instances of absolute systematic behavior. Nevertheless, this number will serve to establish search performance benchmarks.

Lastly, it is not certain that an inspector will detect a particular defect, even though the cell containing the defect has been fixated on. This uncertainty is due to such factors as the conspicuity of the target, and its distance from the center of fixation. The conditional probability that an inspector successfully detects a particular defect, provided that the cell containing the defect has been fixated on, will be referred to as the conditional probability of detection. Two assumptions will be made regarding this conditional probability. First, it does not vary with the cumulative number of fixations, even in the event there are repeated fixations on the same cell. Secondly, it is not affected by the location of a defect in the search field. As a result, for each defect type, t, the conditional probability of detection is considered to be a constant, p_t.[2]

3. Expected search accuracy

Two complementary models of visual search are adopted here for the inspection of aircraft structures. These models were formulated under the exclusive assumptions of systematic and random search behavior in order to encompass the entire range of search performance. The performance measure of interest is accuracy, where accuracy is defined as the proportion of defects that are detected in a particular search field (e.g., an aft cargo pit) within a specified period of time (or equivalently, a given number of fixations). Since accuracy is a random variable, the expected value (mean) of the accuracy was employed as the actual measure of search performance.

Now, suppose that the number of fixations that occur is equal to an integer multiple, m, of the number of distinct cells that comprise the search field, c. For a defect of type t, Harris et al. (1998) have demonstrated that the mean of the search accuracy is

$$\mu_t^r(m \times c) = 1 - [1 - (p_t/c)]^{m \times c}, \quad (1)$$

in the case of strictly random behavior, whereas in the case of strictly systematic behavior the mean is

$$\mu_t^s(m \times c) = 1 - (1 - p_t)^m, \quad (2)$$

for $(m \times c)$ fixations, for $m = 1, 2, \ldots$ and $c = 1, 2, \ldots$ (Note that under the assumption of strictly systematic behavior the search field will be completely scanned m times in $(m \times c)$ fixations, since the number of fixations required to make a complete scan

[2] Certain factors that affect the uncertainty of detection, such as the conspicuity of the target, and its distance from the center of fixation, have been incorporated into the model (refer to Section 4). However, the model does not capture other factors that may also influence the outcome, such as vigilance and expectancy.

corresponds to the number of distinct cells that comprise the search field.)

Then, since it can be shown that

$$\mu_t^r(m \times c) \leqslant \mu_t^s(m \times c) \tag{3}$$

(Harris et al., 1998), it follows that the corresponding range for the mean of the search accuracy, $\mu_t(m \times c)$, is defined by

$$\mu_t^r(m \times c) \leqslant \mu_t(m \times c) \leqslant \mu_t^s(m \times c), \tag{4}$$

since the accuracy yielded by any mixture of random and systematic behavior lies between the two extremes.

As alluded to earlier, accuracy depends on the number of fixations, the conditional probability of detection, and the size of the search field. Specifically, an analysis of Eqs. (1) and (2) reveals that the mean accuracy is directly related to both the number of fixations and the conditional probability of detection, as would be expected, regardless of the type of search behavior. In the case of strictly random behavior, also as anticipated, the mean accuracy is inversely related to the number of cells in the search field. Lastly, it is significant that the accuracy may be obtained without any information pertaining to the number of defects or their location, since this data is normally unavailable in practice.

4. Illustrations and analysis

In this section, the mean search accuracy and its relationship to the various factors that influence it will be explored further. A factor of particular interest here is the search time, due to the importance of the speed/accuracy trade-off. The implications of these findings, as they apply to aircraft inspection, will also be addressed.

It is certainly possible to determine the range of the expected accuracy from the limits for a specific length of time, or equivalently, a certain number of fixations. (A fixation averages about one-third of a second in duration (e.g., Drury, 1985)). Moreover, since the range is defined by the difference between the best-case and worst-case search behaviors, it serves to underscore the relative improvement in accuracy that could be achieved (through training, for example).

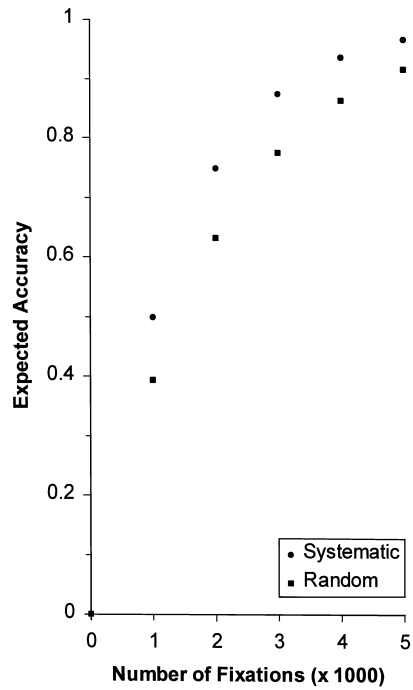

Fig. 1. Expected accuracy versus number of fixations for $c = 1000$ with $p_t = 0.5$.

As an illustration, consider Fig. 1 where the expected accuracies for the cases of strictly systematic and random behavior are depicted. A conditional probability of detection equal to 0.5 and a field size of 1000 cells were chosen for this illustration. (Thus integer multiples of 1000 fixations constitute complete scans for the systematic model.) Observe, for example, that the mean accuracy corresponding to 1000 fixations is bounded below by 0.4 (approx.) and above by 0.5.

The incremental improvements in accuracy that can be attained through the allocation of additional time to inspection can also be determined. Note in Fig. 1, for example, that an additional 1000 fixations will increase the accuracy by about 50%, although this rate of increase declines as the number of fixations increase. This type of information is of course fundamental to assessing the consequences of the trade-off between speed and accuracy.

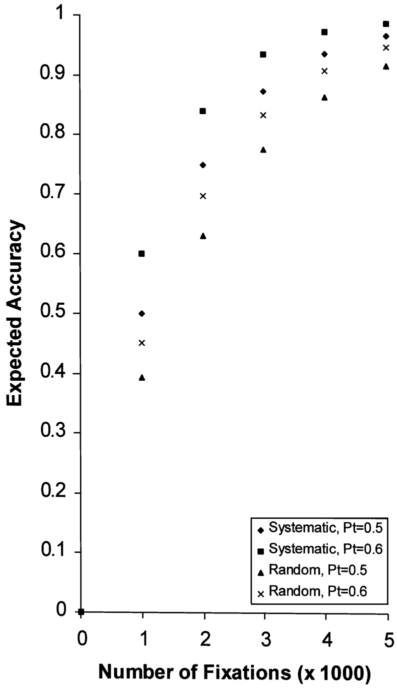

Fig. 2. Expected accuracy versus number of fixations for $c = 1000$ with $p_t = 0.5$ and $p_t = 0.6$.

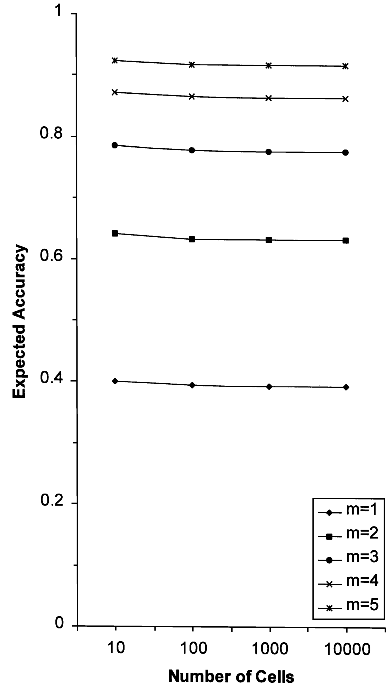

Fig. 3. Expected accuracy versus treatment combinations for random model with $p_t = 0.5$.

In addition to the number of fixations, the expected accuracy will also affected by the conditional probability of detection and the field size. First, recall that the expected accuracy for both models is directly related to the conditional probability of detection. This relationship is confirmed in Fig. 2, where the accuracy limits have shifted upward as a result of increasing the conditional probability of detection to 0.6. A further comparison of Figs. 1 and 2 reveals that the difference between the extremes is larger initially, but gradually diminishes (and eventually becomes smaller) as the number of fixations increases and the expected accuracy approaches its maximum.

Secondly, recall that the expected accuracy for the random model is a function of the field size, whereas the expected accuracy for the systematic model is not. If the proportion of the number of fixations to the number of cells is maintained, however, the expected accuracy of the random model is not sensitive to this parameter. The treatment combinations depicted in Fig. 3 evidence this lack of sensitivity to the relative field size. While these treatment combinations are by no means exhaustive, they are nevertheless representative of situations in which the search fields are composed of more than a small number of cells. The implication of this result is that the expected accuracy will be roughly the same for structures of different sizes, provided that the ratios of the search times to the field sizes are proportional for the various structures.

Lastly, in practice the values for both the conditional probabilities of detection and the number of cells that comprise the search field must be resolved. First, the conditional probabilities of detection can be determined experimentally in any one of several ways (Engel, 1977; Widdel and Kaster, 1981; Bowler, 1990). Engel (1977), for example, describes a procedure that requires a subject to fixate on the center of a region while a target appears at various distances and directions from the fixation point in the periphery. The percentage of the targets that the subject locates, in conjunction with their respective positions, determines the conditional probability of detection. (Note that the value obtained actually represents the *average* conditional probability of detection for a particular defect type.) This procedure can also serve to establish the size of the visual lobe; the ratio of the area of the search field to that of the visual lobe approximates the number of cells that comprise the field.

5. Conclusions

Models for visual search that characterize aircraft inspection tasks were adopted to examine the specific relationship between speed and accuracy in this environment. The two models employed predict the mean accuracy as a function of time under the assumptions of either strictly systematic or random search behavior. In the absence of knowledge of individual search behavior, these two models encompass the entire range of mean accuracy for a given scenario. The magnitude of the difference between the two limiting values indicates the degree of improvement that could be achieved through training. Moreover, the incremental improvements in accuracy that can be attained with increased search time can also be determined.

Subsequently, the relationship of these limits to the other model parameters, the conditional probability of detection, and the size of the search field were examined. Firstly, the mean accuracy is directly related to the conditional probability of detection, as would be expected, regardless of the type of search behavior. Secondly, also as anticipated, the mean accuracy is inversely related to the number of cells in the search field in the case of strictly random behavior. Nevertheless, it was determined that the mean accuracy is not significantly diminished by an increase in the field size if the search time is increased proportionally. Subsequently, a method for determining these parameters in practice was also summarized. Finally, it is noteworthy that the mean accuracy neither depends on the number of defects nor their location, since this data is generally not available in practice.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 9622654. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors, and do not necessarily reflect the views of the National Science Foundation.

References

Arani, T., Karwan, M.H., Drury, C.G., 1984. A variable-memory model of visual search. Human Factors 26, 631–639.

Bobo, S.N., 1990. Communication and transfer of nondestructive inspection information. In: Report of the Second Federal Aviation Administration Meeting on Human Factors Issues in Aircraft Maintenance and Inspection, Office of Aviation Medicine, Federal Aviation Administration, Washington, D.C., pp. 151–166.

Bowler, Y.M., 1990. Towards a simplified model of visual search. In: Brogan, D. (Ed.), Visual Search. Taylor & Francis, London, pp. 3–19.

Drury, C.G., 1985. Stress and quality control inspection. In: Cooper, C.L., Smith, M.J. (Eds.), Job Stress and Blue Collar Work. Wiley, New York, pp. 113–129.

Drury, C.G., 1994. The speed accuracy tradeoff in industry. Ergonomics 37 (4), 747–763.

Drury, C.G., Chi, Chia-Fin, 1995. A test of economic models of stopping policy in visual search. IIE Transactions 27, 382–393.

Drury, C.G., Prabhu, P., Gramopadhye, A., 1990. Task analysis of aircraft inspection activities: methods and findings. In Proceedings of the Human Factors Society 34th Annual Conference, Santa Monica, CA pp. 1181–1185.

Engel, F.L., 1977. Visual conspicuity, visual search, and fixation tendencies of the eye. Vision Research 17, 95–108.

Federal Aviation Administration, 1991. Human Factors in Aviation Maintenance – Phase One Progress Report, Washington, DC, Office of Aviation Medicine, Federal Aviation Administration, DOT/FAA/AM-91/16.

Harris, J.M., Melloy, B.J., Gramopadhye, A.K., 1998. The accuracy of inspectors in time-terminated visual searches. Working Paper 97-102, Department of Industrial Engineering, Clemson University.

Karwan, M.H., Morawski, T.B., Drury, C.G., 1995. Optimum speed of visual inspection using a systematic search strategy. IIE Transactions 27, 291–299.

Morawski, T.B., Drury, C.G., Karwan, M.H., 1980. Predicting search performance for multiple targets. Human Factors 22, 707–718.

Morawski, T.B., Drury, C.G., Karwan, M.H., 1992. The optimum speed of visual Inspection using a random search strategy. IIE Transactions 24 (5), 122–133.

Phillips, E.H., 1994. Focus on accident prevention key to future airline safety. Aviation Week and Space Technology August, 52–53.

Widdel, H., Kaster, J., 1981. Eye movement measurements in the assessment and training of visual performance. In: Moraal, J., Kraiss, K.F. (Eds.), Manned Systems Design, Methods, Equipment and Applications. Plenum Proc, New York, pp. 251–270.

Williams, L.G., 1966. Target conspicuity and visual search. Human Factors 8, 80–92.